Dubovickii-Miljutin Theory and Pontrjagin's Maximum Principle

Kevin R. Vixie

Portland State University and Los Alamos National Laboratory

Abstract

In this talk I give a brief introduction to generalized Lagrange multipliers and then apply them to elucidate the idea of the proof of Pontrjagin's Maximum Principle.

Outline

- Optimal Control: The bang-bang principle
- Constrained Optimization: The usual elementary version
- Constrained Optimization: Dubovickii-Miljutin Theory
- Pontrjagin's Maximum Principle: proof via DM theory
- Example: Classical Mechanics
- References

Optimal Control: The bang-bang principle

Suppose that we are driving a one dimensional vehicle ...

P1.1: Path $a \leq x \leq b$

P1.2: Contraints $\|\ddot{x}\| \leq 1$

P1.3: Initial Conditions $x(0) = a$, $\dot{x}(0) = 0$

P1.4: Final Condition $x = b$

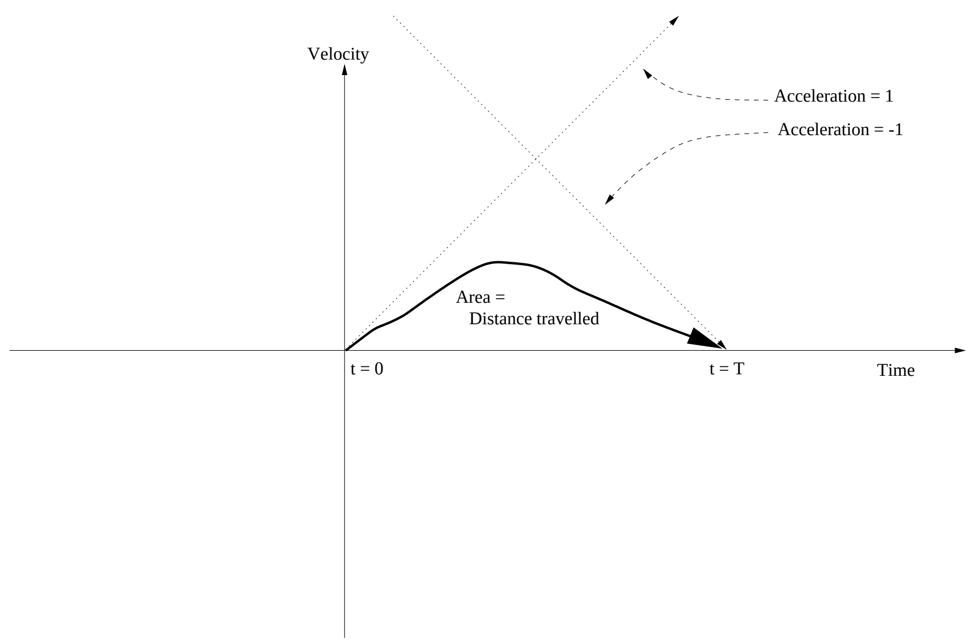
P1.5: Input the acceleration , $\ddot{x}$

P1.6: Objective minimize time taken to go from a to b

Then ...

Theorem. [Bang-Bang Principle] *The solution of P1 is to accelerate at 1 for* $sqrt(4(b-a))/2$ *seconds and then to accelerate at -1 for* $sqrt(4(b-a))/2$ *seconds. (Therefore implying that the optimal input or* control *is discontinuous!)*

A picture ...



Curve lies in trianglular region ... (intermediate value theorem)

Constrained Optimization: The Usual Elementary Case

Typical approach to constrained optimization as seen in a first course in calculus, is to use Lagrange multipliers.

P2.1 Objective Maximize $f(x), f : R^n \to R$

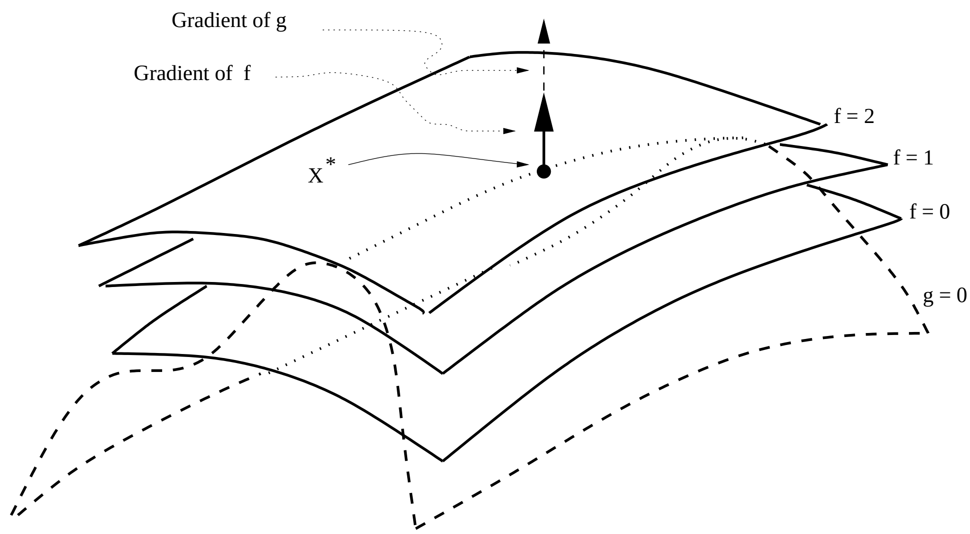
P2.2 Constraint $g(x) = 0, g : R^n \to R$

Theorem. [Necessary Condition] *For x^* to be a solution of P2,*

$$\nabla f(x^*) = \lambda \nabla g(x^*)$$

for some $\lambda \in R$.

Another Picture ...



the normal vectors must colinear ... in fact if we have that $g \leq 0$ as the constraint, then $\lambda \geq 0$.

Constrained Optimization: Dubovickii-Miljutin Theory

Suppose we modify the previous example to a bit more general picture.

P3.1 Objective Maximize $f(x), f : R^n \rightarrow R$

P3.2 Constraints

$G_i(x) \leq 0, i = 1, 2, \ldots, m; G_i : R^n \rightarrow R$

P3.1' Objective Minimize $\acute{f}(x) \equiv -f(x), f : R^n \rightarrow R$

P3.2' Constraints

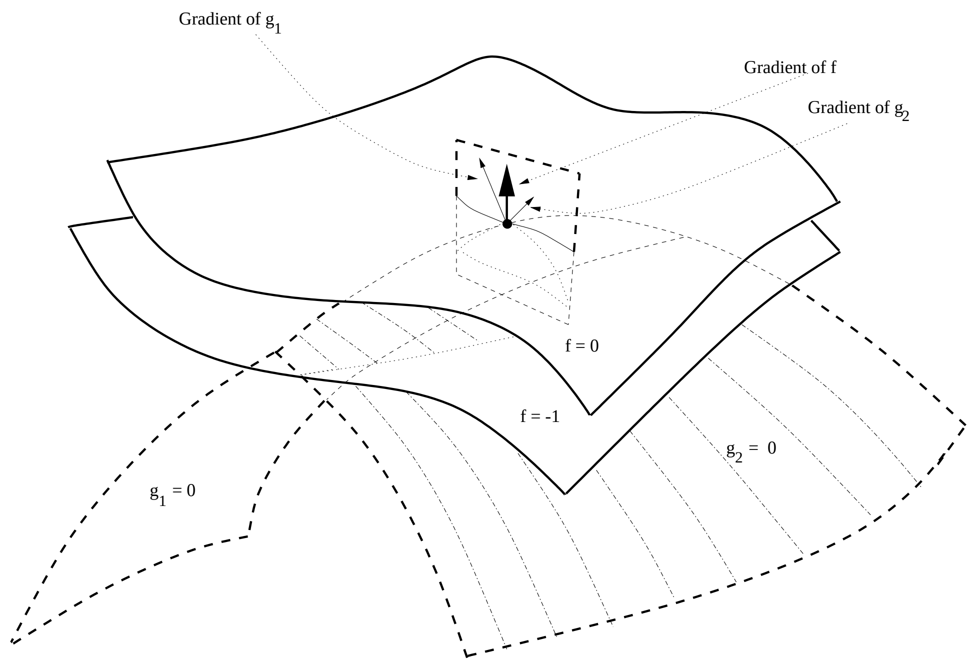
$G_i(x) \leq 0, i = 1, 2, \ldots, m; G_i : R^n \rightarrow R$

Theorem. [Necessary Condition] *For x^* to be a solution of P3,*

$$\nabla f(x^*) = \sum_i \lambda_i \nabla G_i(x^*)$$

for some $\{\lambda_i\}_1^m \in R_+^m$.

Another picture,



In this case we will have that λ_1 and λ_2 are both nonnegative.

Comments

$m < n$ Lagrange relation special
$m = n$ Intersection typically one point
$m > n$ Intersection special ... typically some constraints not active

Cones and Dual Cones

Review ...

Vector Space

Hausdorff

Topological Vector Space

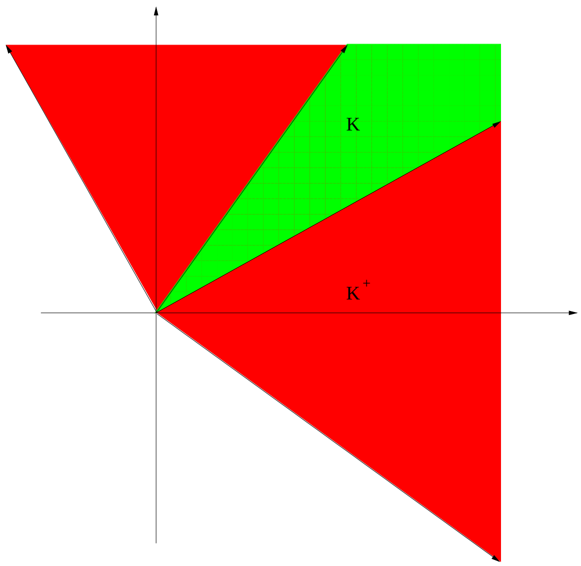
Seminorm

Locally convex spaces

Definition. [Cones and Dual Cones] *1) A subset K, of a locally convex space X is a cone if, given that $k \in K$ and $\lambda > 0$ this implies $\lambda k \in K$. 2) K^+ is the dual cone to K, defined to be all the elements $\hat{k}$ of X^* such that $\hat{k}(k) \geq 0 \; \forall k \in K$.*

For example the dual cone of a half space is a ray normal to that half space. Here we have identified the space and it's dual ...

If we look at a cone K (green) in R^2 and its dual cone K^+ (red) in $(R^2)^*$ we might have something like ...



Facts:

Convexity: Cones need not be convex ... dual cones always are.

Duality $K^{++} = \bar{co}K$ when $K \neq \emptyset$.

Minimum Values If $f \in X^*$ and $K \neq \emptyset$, then $\{\inf f(k)|k \in K\} = 0$ iff $f \in K^+$ and $\{\inf f(k)|k \in K\} = -\infty$ iff $f \ni K^+$.

Now we come to the three main results of this section. Each one is really a special of the previous one.

Suppose that

A1: $K_0, K_1, \ldots, K_{n+1}$ are convex cones in a real LCS X
A2: $K_0, K_1, \ldots, K_n$ are open and $K_0 \neq \emptyset$.

then,

Lemma. [Dubovickii-Miljutin]

$$K_{int} \equiv \bigcap_{i=0}^{n+1} K_i = \emptyset \Longleftrightarrow f_0 + f_1 + \ldots + f_{n+1} = 0$$

where $f_i \in K_i^+$, and f_i are not all zero.

- Examples

Idea of Proof [1, v.3,ch.48]:

($\Leftarrow$)

- assume not ... and $u \in K_{int}$
- w.l.o.g. $f_0(v) \neq 0$ for some v.
- $f_0(u) + \lambda f_0(v) \geq 0$ for all λ in a nbhd. of 0.
- contradiction

($\Rightarrow$)

- (*)
$$\left(\bigcap_{i=0}^{m} K_i \right)^{+} = K_0^{+} + K_1^{+} + \ldots + K_m^{+}$$
for $m \leq n$.
- there is an $m \leq n$ such that
$$K \equiv \bigcap_{i=0}^{m} K_i \neq \emptyset \quad , \quad K_m \cap K_{m+1} = \emptyset$$
- K is open and can be separated from K_{m+1} ... i.e. $\hat{f} \neq 0, \in X^*$ and $\hat{f}(K_{m+1}) \leq a \leq \hat{f}(K)$, $(a \in R)$.

- Therefore, $\hat{f} \in K^+$, $-\hat{f} \in K^+_{m+1}$.
- Use (*) to get that $\hat{f} = f_0 + f_1 + ... + f_m$, some $f_0, f_1, ..., f_m \in (K_0^+, K_1^+, ..., K_m^+)$ and set $f_{m+1} = -\hat{f}$. Set $f_{m+2} = f_{m+3} = ... = f_n = 0$.
- Done (almost)
- idea of proof of (*) still needed.
- ($\supseteq$) easy
- ($\subseteq$) not so easy ... uses Krein extension theorem ...
- $Y \equiv \prod_{i=0}^m X$, it's diagonal L, and $C \equiv \prod_{i=0}^m K_i$.
- $(\prod_{i=0}^m X)^* = \prod_{i=0}^m X^*$.
- $f \in (\bigcap_{i=0}^m K_i)^+$, F defined via f on diagonal.
- $F \geq 0$ on $C \cap L$ ... KE theorem ... $F \geq 0$ on C.
- $F = f_0 + f_1 + ... + f_m \quad , \quad f_i \in X^*$.
- $f_i \in K_i^+$.
- Done!

Second result ... A little more familiar statement!

(remember, we are heading towards constrained max/min problems in Banach spaces!)

Problem: Minimize $F_0(u)$ given,

- Inequality type constraints: $u \in N_j$, $j = 1, ..., n$
- Equality type constraint: $u \in N_{n+1}$

Goal: Find necessary condition of the form

$$f_0 + f_1 + ... + f_{n+1} = 0 \quad , \quad f_i \in K_i^+$$

For some as yet undefined dual cones derived from the constraints.

Cones made up of:

- regular descent directions K_0
- admissable directions K_j , $j = 1, ..., n$
- tangential directions K_{n+1}

Theorem 1. *Given the assumptions,*

- $D(F_0)$ *is a functional in neighborhood of* u_0
- $N_1, N_2, \ldots, N_n$ *are subsets of* X *with nonempty interior, but* N_{n+1} *may have empty interior.*
- $K_0, K_1, \ldots, K_{n+1}$ *are convex and* $K_0 \neq \emptyset$

we have that

Necessary Condition *If* u_0 *is a local solution to the above problem, there exists* f_i*'s in the* K_i^+*'s such that*

$$f_0 + f_1 + \ldots + f_{n+1} = 0$$

.

Nondegeneracy $\bigcap_{i \neq k} K_i \neq \emptyset \quad \Rightarrow f_k \neq 0$

Sufficient Condition *The necessary condition is sufficient to guarentee that* u_0 *is a global minimum if*

- $F_0 : X \to \Re$ *is convex and continuous*
- $N_1, \ldots, N_{n+1}$ *are convex and there is a point* h *in the interior of* $N_1, \ldots, N_n$ *which is also in* N_{n+1}.

Proof of Theorem 1:

... use the Lemma! ... and use some intuition for the Nondegeneracy and Sufficient conditions.

(Theorem 1 is due to Dubovickii and Miljutin (1965))

Third result ... Constrained optimization (infinite dim.)!

As preparation for the next theorem, we consider the problem that will look most familiar to those aquainted with constrained optimization.

Problem: Minimize $F_0(u)$ with the constraints given by

C1 $F_j \leq 0 \ , \ j = 1, ..., n-1$

C2 $u \in N_n$

C3 $F_{n+1}(u) = 0$

... Assuming that we have ...

A1 $F_0, ..., F_{n-1} : X \to \Re$ are F-differentiable functionals.

A2 N_n is a convex subset of X with non-empty interior.

A3 $F_{n+1} : X \to Y$ is a continuously F-differential operator.

A4 X and Y are real Banach spaces.

A5 Regularity: The range $R(F'_{n+1}(u))$ is closed in Y.

Then we would like nec. and suf. conditions ...

Theorem 2. [Generalized Kuhn-Tucker Theory] *For the above minimization problem we have the following condition that is necessary for u_0 to be a local solution.*

$$\sum_{i=0}^{i=n-1} \lambda_i DF_i(u_0)(u-u_0)+\langle y^*, DF_{n+1}(u_0)(u-u_0)\rangle \geq 0 \;\; \forall u \in N_n$$

where

$$\textbf{\textit{multipliers:}} \; \lambda_i \geq 0 \quad , y^* \in Y^*$$
$$\textbf{\textit{non-degeneracy condition:}} \; \lambda_i F_i(u_0) = 0$$

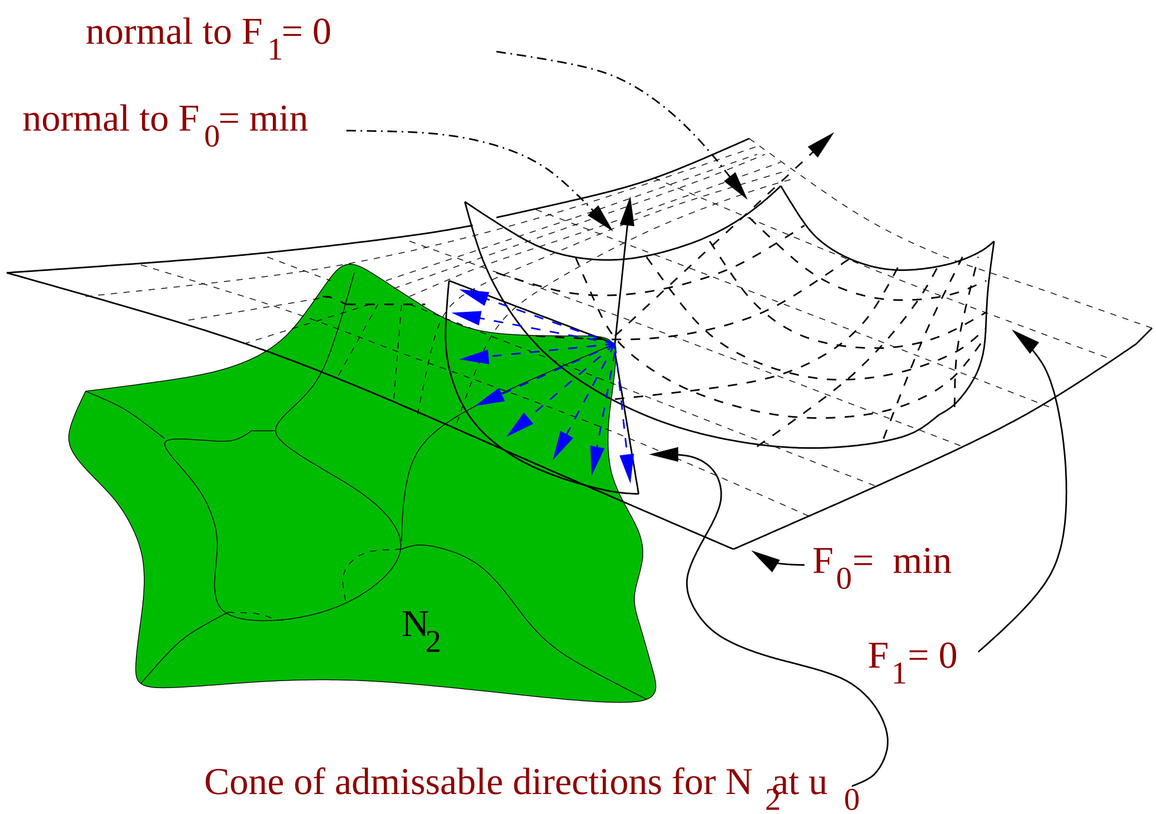
for $i = 0, \ldots, n-1$.

If $\lambda_0 > 0$ and $F_0, \ldots, F_{n-1}, \langle y^*, F_{n+1}(\cdot)\rangle$ are convex on X, then the condition is sufficient for u_0 to be a global solution to the minimum problem.

A Pause to Collect and Consider ...

- Theorem 2 "=" Theorem 1
- Theorem 2 permits us to work in functional spaces!!

Another picture ...



... on to Pontrjagin's maximum principle.

Pontrjagin's Maximum Principle: proof via DM theory

The maximum principle concerns itself with solutions to a control problem that minimize some functional. For example, one might want to minimize the time taken to move from point A to point B under constrained controls. This is what we considered in the bang-bang principle above.

In the following we will focus on the statement of the principle and a very brief overview of how the proof is approached. The details will be completely glossed over with a couple of exceptions. For the details, one should consult [1, v.3,ch.48,p.422-33]. What I try to clarify is the fact that we prove the PMP by pushing the problem into a form permitting the application of Theorem 2 ... modulo the introduction of a time scaling function ... and then get the result via the introduction of adjoint states.

Optimal Control Problem P:

Control Functional: $\int_{t_1}^{t_2} f(y(t), w(t), t)dt \quad = min!$
Control Equations: $\dot{y}_i(t) = g_i(y(t), w(t), t)$
Boundary conditions: $h_i(t_2, y(t_2)) = 0$
Control constraints: $w(t) \in W^{[t_1,t_2]}$
...where: $i = 1, ..., N$; $y_i(t_1) = a_i$; f, g_i, and h_i are C^1

Comments:

- This covers a huge amount of ground ... note that we obtain discontinuous vector fields through the dependance upon the control $w(t)$ which can be discontinuous.
- we assume t_1 is fixed, but that t_2 is determined by the solution.

Define:

$$H(y, w, p, t, \lambda_0) \equiv \sum_{i=1}^{N} p_i g_i(y, w, t) - \lambda_0 f(y, w, t)$$

The Theorem ... at last!

Theorem 3. [Pontrjagin's Maximum Principle] *Given that $(y(t), w(t), t_2)$ is a solution to P we have that,*

- $\exists \lambda_0, \alpha_1, ..., \alpha_N$, *not all zero and* $\lambda_0 \geq 0$
- $\exists$ *functions* $p_i(t)$, $i = 1, ..., N$, *continuous on* $[t_1, t_2]$

such that

$$H(y(t), w(t), p(t), t, \lambda_0) = \max_{w^* \in W} H(y(t), w^*, p(t), t, \lambda_0))$$

and

$$\dot{p}_i = -H_{y_i}\ , \qquad y_i = H_{p_i}\ , \qquad i = 1, ..., N$$

$$p_i(t_2) = -\sum_{j=1}^{N} \frac{\partial h_j}{y_i}(t_2, y(t_2))\alpha_j\ , \qquad i = 1, ..., N$$

Idea of Proof:

- introduce $t = v(\tau)$ to transform time t to $\tau \in$ [0,1].
- compute the derivatives
- $\lambda_0 F_0^{'}(\bar{u})(u - \bar{u}) + \langle y^*, F_2^{'}(\bar{u})(u - \bar{u}) \rangle \geq 0$
- introduce adjoint coordinates ... (looks right now)
- switch back to time t.
- done

Next an example ... the Euler-Lagrange equation of classical mechanics.

Example: Classical Mechanics

- $\int_0^T L(x(t), u(t))dt = \min!$
- $\dot{x}(t) = v(x(t), u(t))$, on $[0, T]$
- T fixed
- $u(t) \in U$

Define $H(x, u, p) \equiv pv(x, u) - L(x, u)$ and we get

- $\dot{p}(t) = -H_x(x(t), u(t), p(t))$
- $p(T) = 0$
- $H(x(t), u(t), p(t)) = \max_{u \in U} H(x(t), u, p(t))$
- ... which is $H_u(x(t), u(t), p(t)) = 0$ for smooth H.

SO, specializing to the case where $v \equiv u$ and $U = \Re$, we get

- $u(t) = \dot{x}(t)$
- $H = p(t)\dot{x}(t) - L(x(t), \dot{x}(t))$
- $p(t) = L_{\dot{x}}(x(t), \dot{x}(t))$
- $\dot{p}(t) = \frac{d}{dt}L_{\dot{x}}(x(t), \dot{x}(t) = L_x(x(t), \dot{x}(t))$

Summary

Outline reiterated:

- Optimal Control: The bang-bang principle
- Constrained Optimization: The usual elementary version
- Constrained Optimization: Dubovickii-Miljutin Theory
- Pontrjagin's Maximum Principle: proof via DM theory
- Example: Classical Mechanics

The moral of the story is that Pontrjagin's Maximum Principle boils down to a result about cones and dual cones in Banach spaces ... that at optimal points the intersection of admissable cones is empty and that this can be translated into a generalized langrange multiplier sum of dual vectors that is non-negative in some cases and identically zero in others. (admissable cones = cones of admissible directions)

Reiteration of key result ...

Key Result:

Suppose that

A1: $K_0, K_1, \ldots, K_{n+1}$ are convex cones in a real LCS X

A2: $K_0, K_1, \ldots, K_n$ are open and $K_0 \neq \emptyset$.

then,

Lemma. [Dubovickii-Miljutin]

$$K_{int} \equiv \bigcap_{i=0}^{n+1} K_i = \emptyset \Longleftrightarrow f_0 + f_1 + \ldots + f_{n+1} = 0$$

where $f_i \in K_i^+$*, and* f_i *are not all zero.*

References

[1] Eberhard Zeidler. *Nonlinear Functional Analysis and its Applications*. Springer Verlag, 1985. (in five volumes).